

Chimenea and other tools: Automated imaging of multi-epoch radio-synthesis data with CASA [☆]

Tim D. Staley^a, Gemma E. Anderson^a

^a*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK*

Abstract

In preparing the way for the Square Kilometre Array and its pathfinders, there is a pressing need to begin probing the transient sky in a fully robotic fashion using the current generation of radio telescopes. Effective exploitation of such surveys requires a largely automated data-reduction process. This paper introduces an end-to-end automated reduction pipeline, **AMIsurvey**, used for calibrating and imaging data from the Arcminute Microkelvin Imager Large Array. **AMIsurvey** makes use of several component libraries which have been packaged separately for open-source release. The most scientifically significant of these is **chimenea**, which implements a telescope-agnostic algorithm for automated imaging of pre-calibrated multi-epoch radio-synthesis data, of the sort typically acquired for transient surveys or follow-up. The algorithm aims to improve upon standard imaging pipelines by utilizing iterative RMS-estimation and automated source-detection to avoid so called ‘Clean-bias’, and makes use of CASA subroutines for the underlying image-synthesis operations. At a lower level, **AMIsurvey** relies upon two libraries, **drive-ami** and **drive-casa**, built to allow use of mature radio-astronomy software packages from within Python scripts. While targeted at automated imaging, the **drive-casa** interface can also be used to automate interaction with *any* of the CASA subroutines from a generic Python process. Additionally, these packages may be of wider technical interest beyond radio-astronomy, since they demonstrate use of the Python library **pexpect** to emulate terminal interaction with an external process. This approach allows for rapid development of a Python interface to any legacy or externally-maintained pipeline which accepts command-line input, without requiring alterations to the original code.

Keywords: methods: data analysis, techniques: interferometric

1. Introduction

The science of transient astronomical events is a burgeoning sub-field of astronomy. For the past decade, optical and gamma-ray transient surveys have led the way, with radio-band observations being largely restricted to follow-up of manually selected targets discovered at other wavelengths. However, a new generation of radio facilities composed primarily of Square Kilometre Array (SKA) pathfinders such as ASKAP (the Australian Square Kilometre Array Pathfinder; Murphy et al., 2013), MeerKAT (Karoo Array Telescope; Booth and Jonas, 2012), LOFAR (the Low Frequency Array; van Haarlem et al., 2013), MWA (Murchison Widefield Array; Tingay et al., 2013; Bell et al., 2014), and the Jansky Very Large Array (JVLA), are now at various stages of commissioning and planning with some (LOFAR, JVLA, MWA) already producing new surveys and the rest scheduled for commissioning within the next few years (see e.g. Norris et al., 2013, for an overview). These new observatories, and in time the SKA, will provide wider fields of view, greater sensitivity and potentially much

faster response to targets of opportunity than was previously possible. This makes them attractive instruments for performing transient science, both for routine multi-wavelength follow-up and a new generation of dedicated surveys in the radio-band (Stappers, 2013). Estimates of transient-discovery rates from SKA surveys currently vary quite widely (see e.g. Burlon et al., 2015; Metzger et al., 2015, for orphan gamma-ray burst rate estimates that differ by an order of magnitude, primarily due to different choices of signal-to-noise thresholds), but it is safe to assume that the total number of transient alerts from across the spectrum will only increase over the coming years, primarily due to optical survey projects such as the Large Synoptic Survey Telescope (Ridgway et al., 2014). Exploiting the potential of the the SKA and its pathfinders for transient science requires that we follow-up these transient alerts in much greater numbers compared to current practice. As such there is a pressing need to begin probing the transient sky in a fully-robotic fashion using the current generation of radio telescopes, in order to develop the techniques and infrastructure that will be required.

Since 2012 we have been using the Arcminute Microkelvin Imager Large Array (AMI-LA, Zwart et al., 2008) to conduct a programme of automated radio-followup of

[☆]Generated from git source 6b20b9d dated 2015-08-25. Source available at github.com/timstaley/automated-radio-imaging-paper
Email address: tim.staley@physics.ox.ac.uk (Tim D. Staley)

transients detected at other wavelengths, now known as the AMI-LA Rapid Response Mode (ALARRM). ALARRM has produced some of the earliest gamma-ray burst (GRB) follow-up in the radio (Staley et al., 2013; Anderson et al., 2014), and unprecedented early-time coverage of a radio-flare from an M-dwarf star accompanying a hard X-ray outburst (Fender et al., 2015).

Existing facilities such as AMI-LA typically rely upon mature data-reduction pipelines created with an interactive, user-intensive model in mind, making them ill-suited for integration into fully-automated analyses of the sort needed for a quick evaluation of transient candidates. However, these legacy pipelines are the result of many developer-years of effort in terms of algorithm development, software engineering, and testing. As such they represent an irreplaceable part of the data-reduction process for existing observatories.

In the course of the ALARRM programme we have developed modular software, written in Python, to automate the data-reduction process. We have recently open-sourced this software as four separate packages.¹ The first two, **drive-ami** and **drive-casa** (Staley and Anderson, 2015a; Staley, 2015), are libraries developed to provide interfaces to mature radio-astronomy software packages, to enable their use from within Python scripts. A third library, **chimenea** (Staley and Anderson, 2015b), implements an heuristic algorithm for automated imaging of pre-calibrated multi-epoch radio-synthesis data, using **drive-casa** to build upon subroutines from CASA (the Common Astronomy Software Applications package, McMullin et al., 2007; CASA Consortium, 2011). The final package, **AMIsurvey** (Staley and Anderson, 2015a), ties together the previous codes to create a telescope-specific end-to-end pipeline, which ingests raw AMI-LA datafiles and produces calibrated and cleaned multi-epoch images ready for source extraction and transient identification.

This paper is structured as follows: In Section 2 we introduce AMI-LA, and describe a new fully scriptable Python interface to its associated calibration pipeline. Section 3 discusses our choice to build upon the CASA routines to implement the imaging stage, and explains the rationale and design of our external scripting interface, **drive-casa**. Section 4 describes the automated imaging algorithm implemented in the **chimenea** package, and Section 5 gives an overview of the end-to-end ALARRM data-reduction pipeline encoded by **AMIsurvey**. Section 6 presents a basic performance analysis demonstrating that the pipeline behaves as expected. In the discussion section, we describe known limitations of our **chimenea** algorithm, and cover some possible extensions and alternatives (§7.1). We then describe and compare different methods for solving a problem relevant in the wider context of software in academia

— that of interfacing with legacy software — and explain the choice of method for this work (§7.2). We finish the discussion with some general recommendations for writing re-usable Python codes (§7.3), and conclude in Section 8.

2. Automated calibration of AMI-LA data with **drive-ami**

The Arcminute Microkelvin Imager Large Array (AMI-LA) is a synthesis telescope composed of eight equatorially mounted 12.8 m dishes sited at the Mullard Radio Astronomy Observatory (MRAO) at Lord’s Bridge, Cambridge. The telescope observes in the band 12.5–18.2 GHz with eight 0.72 GHz bandwidth channels.

AMI-LA observations are recorded in a custom data format to be processed by a specialized pipeline, **AMI-REDUCE**. This applies path delay corrections, automatic flags for interference, pointing errors, shadowing and hardware faults, applies phase and amplitude calibrations, Fourier transforms the data into the frequency domain, and writes out the resulting data in *uv*-FITS format (Davies et al., 2009). Typically data reduction is either carried out interactively, allowing the user to view plots of the data in order to guide the reduction process, or by calling one of a selection of standard reduction scripts (which are simple listings of **AMI-REDUCE** commands, without any flow control such as ‘if’ statements or ‘for’ loops.).

Application of one of the standard **AMI-REDUCE** listings was adopted as a calibration step for observations made as part of the ALARRM programme. Early versions of the ALARRM pipeline ran **AMI-REDUCE** from a Python script via a simple subprocess call (cf §7.2.2), but this proved problematic. First, when inspecting the logs, a successful reduction with **AMI-REDUCE** was indistinguishable from an aborted reduction due to a problematic dataset. Second, important metadata such as the rain modulation (which gives some idea of the data-quality) and observation time is not stored in the header of the *uv*-FITS output files, and so checking these values still required manually loading the dataset in question into the **AMI-REDUCE** pipeline.

Clearly, a more powerful programmatic interface to **AMI-REDUCE** was required. Accordingly, we implemented an interface that provides a set of Python routines for controlling **AMI-REDUCE**, under the moniker **drive-ami**. To do so we make use of **pexpect**², a pure-Python library for emulating terminal-interaction, as discussed in subsections 7.2.3 and 7.2.4.

In addition to allowing step-by-step calling of the **AMI-REDUCE** functionality from a Python script, **drive-ami** provides a number of enhancements to the data-reduction process. Data-provenance and reproducibility are improved; all the **AMI-REDUCE** commands are recorded in a log-file along with the output *uv*-FITS data. In a separate file,

¹The packages described in this paper are available from <https://github.com/timstaley/drive-ami>, <https://github.com/timstaley/drive-casa>, <https://github.com/timstaley/chimenea>, and <https://github.com/timstaley/amisurvey>.

²<http://pexpect.readthedocs.org>

all textual output displayed during the reduction process is logged alongside the input commands, so that the end-user may read back over the reduction dialogue as if scrolling back through a terminal history arising from using AMI-REDUCE interactively. A variety of metadata such as observation times, calibrator source, rain modulation and percent of data flagged due to radio-frequency interference are also parsed from the output and stored in JavaScript Object Notation (JSON) format, which has the benefit of being both human-readable and trivially parsed in most high-level programming languages.

As an added convenience, `drive-ami` can also be used to help configure reduction of multiple-epoch data for a source observed with AMI-LA. While observations are usually easy to sort by their filename-prefix (which refers to the target name), occasionally a target is renamed after the initial epoch of observation (e.g. to reflect an assigned GRB ID), or a prefix may be incremented if multiple observations are performed in a single day. `drive-ami` can utilize AMI-REDUCE to extract the pointing information from the raw data-files and then group observations according to their pointing, given a tolerance limit on the maximum angular separation between pointing-centres. We make use of the `astropy` co-ordinate routines for calculating the angular separation (Astropy Collaboration et al., 2013). The JSON-encoded metadata accompanying the resulting *uv*-FITS data then contains a ‘group’ tag, which can be used to identify all observations of a given target for processing as a group (cf Section 4).

3. External scripting of CASA subroutines with `drive-casa`

Traditionally, the *uv*-FITS data products output from the AMI-REDUCE reduction stage are reduced using the AIPS package (Fomalont, 1981), but for the ALARM programme we chose to make use of CASA (McMullin et al., 2007). We made this switch for a few reasons. We judge CASA to be the best supported general-purpose package for radio-astronomy data-reduction, with a simple installation process, frequent updates³, a comprehensive accompanying manual⁴, and a helpdesk service. CASA is also widely applicable, recognising data from many of the current generation of telescopes; support is also available for newly commissioned facilities such as LOFAR⁵ and ALMA. Lastly, the Python interpreter underlying the default CASA interface, `casapy`, provides a convenient scripting interface for many basic tasks, albeit with some restrictions, as described below.

It is trivial to run simple Python scripts within the `casapy` environment, or even to launch `casapy` into a

Python script directly from the command line. However, problems arise if you wish to make use of functionality from both `casapy` and non-standard Python packages (i.e., anything except the standard library) from within the same script. This is because `casapy` bundles its own Python interpreter, customized to display a logging window and provide access to plotting tools. As a result, Python packages installed in the normal manner are unavailable from within the `casapy` environment. This is unfortunate, as it precludes easy usage of CASA routines from within a larger pipeline, even though much of the underlying interface is already written in Python.

With regards to data-provenance and reproducibility, `casapy` provides extensive logfiles which provide a record of the commands and resulting information from a data-reduction run. However, a minor flaw is that once instantiated, `casapy` provides no method (or at least, no documented method) for redirecting log output, which can make it hard to locate the log section relevant to a given observation when reduced as part of a batch. Logging of input and output is also collapsed into a single stream, which makes it tricky to extract the commands required to reproduce a reduction process.

Programmatic error-handling is also lacking from the `casapy` environment. The standard `casapy` routines do not provide a return value, or throw exceptions on encountering error conditions. Instead, all processing information and errors are output as log messages, which makes it hard to respond in an automated fashion.

Our package for automating interaction with `casapy`, ‘`drive-casa`’, addresses these issues. Similarly to `drive-ami`, we use `pexpect` to emulate terminal interaction with the `casapy` process, as detailed in Section 7.2.3.

The `drive-casa` interface can be used to execute strings or scripts containing `casapy` commands directly, so all `casapy` subroutines are accessible via the `drive-casa` interface. Alternatively, the package includes a set of convenience routines which try to adhere to a consistent style and make it easy to chain together successive CASA reduction commands to generate a `casapy` script programmatically. For example, the following code fragment generates a script to invoke the CASA routine `importUVFITS`, then perform a zero-iteration Clean operation on the resulting CASA *MeasurementSet* to produce a dirty map:

```
script = []
ms = drivecasa.commands.import_uvfits(
    script,
    uvfits_path)
dirty_maps = drivecasa.commands.clean(
    script,
    ms,
    niter=0,
    threshold_in_jy=1,
    other_clean_args=clean_args)
```

³http://casa.nrao.edu/previous_casa.shtml

⁴https://casa.nrao.edu/ref_cookbook.shtml

⁵See also the LOFAR imaging cookbook:

<http://www.astron.nl/radio-observatory/lofar/lofar-imaging-cookbook>

chimenea:
An heuristic algorithm for automated imaging
of pre-calibrated multi-epoch radio-synthesis data

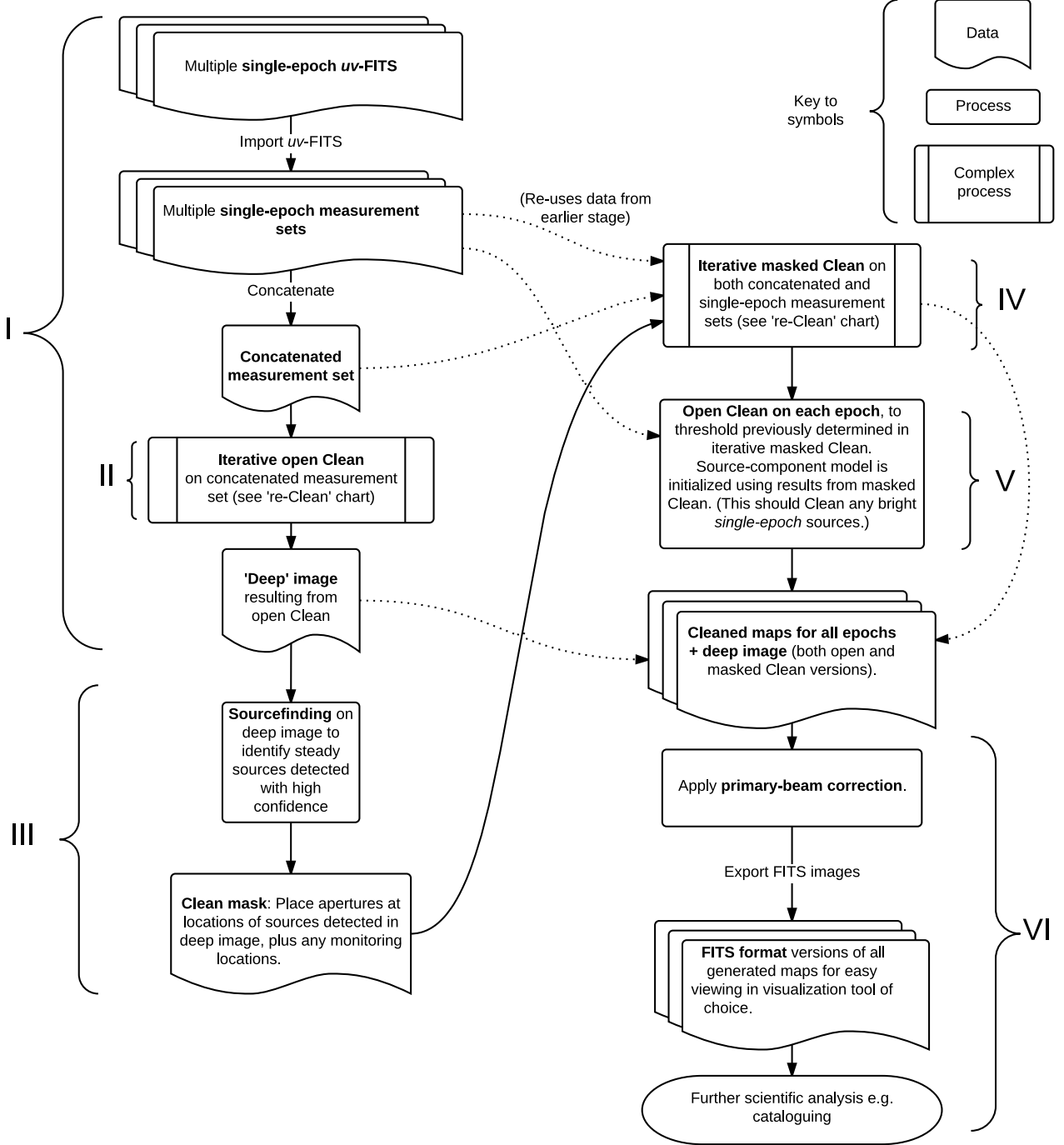


Figure 1: An overview of the algorithm implemented in the **chimenea** pipeline. Dotted lines denote re-use of data-products from an earlier stage. Groupings labelled in Roman numerals refer to descriptions in text. The terms ‘open’ and ‘masked’ Clean refer to applying the Clean algorithm (Schwab, 1984) either to a box-region covering the central quarter of a map (i.e. an open or unconstrained clean), or to a masked map, i.e. one where the Clean algorithm is constrained to placing model components in the vicinity of known sources. The ‘re-Clean’ subroutine is depicted in Figure 2.

We note that the recently released package `casa-python`⁶ should make it easier to install external Python packages into the `casapy` environment, which goes some way to addressing the issue of package interoperability whilst retaining access to the standard `casapy` logging interface and plotting tools. This may provide a preferable solution for those who wish to make small alterations to their `casapy` workflow to include external functionality, but still effectively treats `casapy` as the top-level pipeline framework, rather than abstracting it to a more standard Python package, as `drive-casa` does.

Since `casapy` is largely Python code internally, and is still undergoing active development, we hope that eventually a native Python interface to CASA *as a library* (not an executable) will be made available, thus negating the need for the terminal-emulation layer of `drive-casa`. In the meantime, `drive-casa` provides a reliable method for incorporating CASA functionality into a fully automated Python reduction pipeline, as demonstrated in Section 4.

4. Automated imaging of multi-epoch radio-synthesis data with `chimenea`

Prompt analysis of results from the ALARRM programme requires frequent re-reduction of target-datasets as additional observations become available. In order to minimise the human effort associated with the imaging stage of the data-reduction, we created a tool for automated imaging of pre-calibrated multi-epoch radio-synthesis data, ‘`chimenea`’, which is a small pipeline built atop the CASA subroutines (using `drive-casa` to provide an interface layer).

The `chimenea` pipeline is designed to automatically choose a suitable set of parameters for applying the Clean algorithm (Högbom, 1974; Schwab, 1984), specifically determining two key values: how ‘deep’ should we Clean (i.e., at what level of peak pixel value in the residuals map do we terminate the Clean process), and how the placing of model components should be constrained via the Clean mask. When performing a blind search for transients, unconstrained or ‘open’ Clean reductions are preferable, since we do not know where a transient source may appear, and an ‘un-Cleaned’ source will be harder to detect since the flux remains spread over the core and sidelobes of the dirty beam. As such, when imaging a field of view which contains no detectable steady sources (to search for transient sources), a simple approach to automating this process is to estimate the root-mean-square noise-level (RMS) from the dirty map, and then apply an open-Clean process, with the Clean termination threshold set to a multiple of the estimated RMS (typically around three).⁷ However, this simple approach produces poor results when a bright steady

source is located in the field of view, particularly if the single-epoch data being imaged is from a short integration with poor *uv*-plane coverage. In this case, the RMS as estimated from the dirty map will often be positively biased, due to additional pixel-value variance from the side-lobes of the bright source. Additionally, an open-Clean process may erroneously place model components in the side-lobes of the bright source, leading to artificial source-artefacts and reduced flux assigned to the model component at the true location of the bright source (a phenomenon known as ‘clean bias’, see e.g. Condon et al., 1998; White et al., 1997, for details). Therefore, when imaging a field containing a known bright source a better strategy is to perform a constrained (‘masked’) Clean — but this potentially leaves transient sources ‘un-Cleaned’, as noted previously.

Previous radio-transient surveys have tackled this problem in different ways. Bell et al. (2014) performed phase and amplitude self-calibration on any standard calibration sources in the field of view, modelled and subtracted the contribution to the visibilities from those calibrators, and then finally applied an open-Clean process to the resulting image with a very limited number of Clean iterations, to avoid over-cleaning any remaining sources. This is effective at reducing side-lobes from calibration sources but will produce sub-optimal results if a bright source is present which is not identified as a calibrator. Alternatively, a manual round of source-identification may be employed prior to performing a final masked-Clean (Miller et al., 2008; Mooley et al., 2013), which is undoubtedly effective but can quickly become labour-intensive in the case of transient surveys, and is unsuitable for near-real-time analysis in any case.

The algorithm encoded in the `chimenea` package takes an iterative approach that makes use of all available epochs of data for a given survey field. Initially, the data from all epochs are concatenated and an open-Clean process is applied, which is fairly robust to artefacts given enough epochs of data, due to improved *uv*-plane coverage. To ensure the RMS-estimation (and hence Clean-termination threshold) is not biased by side-lobes in the dirty map, we re-estimate the RMS after applying an open Clean and then reapply the Clean process with a lower threshold if a significant drop in RMS is detected. The deep image synthesized from the concatenated data is then used to search for any steady sources in the field of view which are detectable with a high signal-to-noise ratio. Next, each single-epoch observation is processed individually. We initially apply a masked Clean, to constrain the placement of Clean-model components to the location of any known bright sources in the field (as detected in the deep image) — this avoids the issues of side-lobe artefacts due to poor *uv*-coverage in shorter observations. Finally, we apply an open

⁶<https://github.com/radio-astro-tools/casa-python>

⁷Alternatively, if the performance of the telescope is well characterised, the visibilities may be analyzed directly to determine the total integration time excluding flagged data, and this value may be

used to directly estimate the RMS noise level. However, ‘well characterised’ can be a difficult ideal to attain when telescope sensitivity may depend on myriad factors such as atmospheric moisture level, local ambient temperature, etc, and even then image-RMS may differ depending on proximity to a bright source.

Clean to each epoch (initialized using the source-model components from the masked clean) in case a transient source is present at detectable flux-density levels.

Figure 1 provides an overview of the implemented data-flow. A step-by-step description of the process, relating to the different stages denoted in Figure 1, is as follows:

- I Concatenating data from all epochs gives better uv -coverage and produces a ‘deep’ map with lower noise, which is ideal for reliably locating steady sources in the field of view. To generate this deep map we apply the Cotton-Schwab variant of the Clean algorithm (Schwab, 1984) to the concatenated data, allowing unconstrained placing of model components within a box-region corresponding to the central quarter of the synthesized map (referred to as an ‘open Clean’ in Figure 1). Sources outside this central region are attenuated by a primary beam response factor of less than 1 in 20,000 relative to the pointing centre, and so a wider open-Clean region is unnecessary unless this margin happens to contain a source with flux level on the order of one Jansky.
- II Applying the Clean algorithm requires that we give a pixel-value termination threshold, which is typically chosen as a multiple of the background root-mean-square (RMS) noise level. However, in order to estimate the RMS, we require an image to sample. To overcome this ‘chicken-and-egg’ problem we take an iterative approach (‘re-Clean’), as depicted in Figure 2. Beginning with the dirty map, we take the pixel values from the residuals map (i.e. the image with any Clean-model components subtracted), apply sigma-clipping to estimate the background RMS, apply Clean with the resulting calculated threshold, and then iterate by examining the new residuals map. Our chosen convergence criteria for ending the process is to stop when the proportional decrease in RMS estimates between iterations is less than a user-chosen value (typically 5%) or when a set number of iterations have completed (typically three), whichever comes first. For the sigma-clipping and RMS estimation step we employ the algorithm described in Section 4.3.1 of Swinbank et al. (2015), which uses information on the synthesized beam shape to correct for bias in the RMS estimate caused by inter-pixel correlation.
- III Applying an initially unconstrained iterative re-Clean process to each single-epoch observation can give poor results, especially if bright steady sources are present in the field of view, as poor uv -coverage can lead to misidentification of side-lobes as separate sources, and over-estimation of background RMS levels. However, since the deep image is generated from multiple concatenated observations it typically has much better uv -coverage and sensitivity. Therefore, we use the deep image to identify steady sources and generate a Clean mask. This mask is used when imaging single-epoch data to constrain the Clean process to place model components only in the vicinity of known sources. To identify sources in the deep image we make use of a source-extraction tool tailored to radio-data and currently distributed as part of the LOFAR transients pipeline (TraP contributors, 2014; Swinbank et al., 2015). The user can also specify co-ordinates of known sources to create additional mask apertures, in order to ensure proper cleaning of any known faint sources or transient candidates.
- IV Applying the constrained (via use of the Clean aperture-mask) re-Clean algorithm to the single-epoch observations is generally quite robust and provides a good estimate of the background RMS at each epoch, in addition to accurately cleaning sources previously detected in the deep image.
- V We can now apply an unconstrained Clean operation to each single-epoch observation, down to the flux-threshold determined by the previous masked re-Clean process, in case there are single-epoch transient sources present which do not correspond to a mask aperture. We initialize this Clean operation with the source-component model derived from the masked Clean process, which should help to ensure that any bright sources in the field are cleaned correctly, without spurious assignment of source-components to bright side-lobe artefacts.
- VI Finally, the pipeline applies primary-beam corrections, and outputs images in both CASA *MeasurementSet* and FITS formats.

We note that the data-flow currently implemented in **chimenea** has been guided by our experience in reducing the ALARRM datasets, and may change in the future as we gain further experience. The current pipeline is already a considerable improvement (see Section 6) over our initial automated solution, which was to simply perform an independent open Clean on each single-epoch observation (Staley et al., 2013), but particularly challenging or scientifically significant datasets will still warrant manual care and attention; **chimenea** provides a reduction which can sometimes be improved further but is good enough for scientific analysis in most cases. Additionally, **chimenea** represents significant value as a working software-implementation, complete with extensive logging, a clearly structured logic-flow, and a number of simple data-structures which aim to make tractable the problem of managing numerous intermediate datasets and configuration parameters. We hope that others will be able to use and build upon this initial implementation.

To encourage re-use, the **chimenea** package has been kept as simple as possible, consisting solely of some basic data structures and the subroutines that make up the algorithm described above, with the complexities of calling out to **casapy** delegated to the **drive-casa** layer. However,

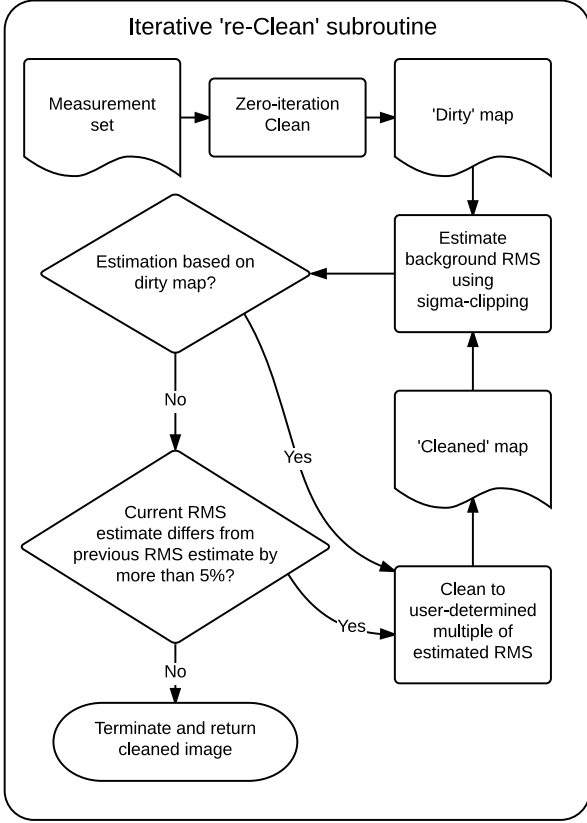


Figure 2: A schematic of the iterative ‘re-Clean’ subroutine, applied as part of the **chimenea** algorithm depicted in Figure 1. The subroutine repeatedly applies the Clean algorithm until acceptance criteria are met. The convergence tolerance (shown as 5% in figure), and the Clean threshold as a multiple of the estimated root-mean-square background noise level (RMS), are both user-configurable. An additional termination criterion is applied — maximum number of iterations around the Clean / re-estimate RMS loop (typically three) — but is omitted from the chart for brevity.

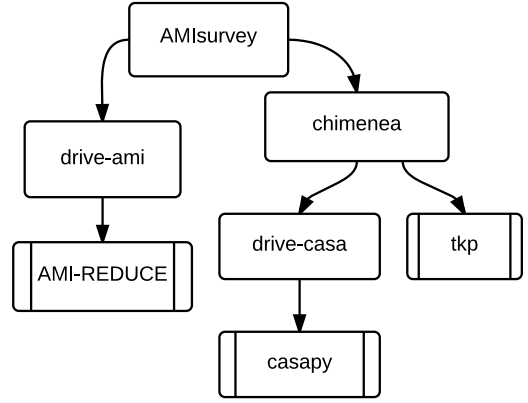


Figure 3: A dependency chart for the **AMIsurvey** package; **AMIsurvey** depends on **drive-ami** which in turn depends on the external package **AMI-REDUCE**, etc. The **tkp** block refers to the LOFAR Transients Key Project software package (TraP contributors, 2014) described in Swinbank et al. (2015). Note that only the salient dependencies are listed; in reality we additionally make use of a number of external Python packages.

this minimalist approach means that there is no command-line interface or default parameter set. These needs are met by the **AMIsurvey** package, described below.

5. **AMIsurvey**, an end-to-end pipeline for automated reduction of AMI-LA observations

The top-level tool used for data-reduction in the ALARRM programme is **AMIsurvey**. This package glues together the calibration and imaging processes implemented via **drive-ami** and **chimenea**, depending in turn upon the **CASA** and **AMI-REDUCE** packages (cf Figure 3). We make use of the the Python standard library package **argparse** to provide a user-friendly command line interface. Initially, the **AMIsurvey** tool loads lists of calibrated *uv*-FITS files and associated metadata from the JSON files output by **drive-ami**. The data-locations and metadata are then trivially converted into the data-structure used by the **chimenea** pipeline, before being sorted into groups. These calibrated observations may be filtered according to quality-control criteria (e.g. limits on rain amplitude modulation) to identify and remove any problematic observations. The grouped-and-filtered observations are then passed, along with configuration parameters (e.g. Clean image-synthesis settings) appropriate to the ALARRM data, to the main **chimenea** pipeline function for processing. Finally, a full listing of the data-products is output in JSON format, for easy ingestion by further analysis tools.

6. Performance testing and results

To thoroughly test the **AMISurvey** pipeline, we reduced a significant portion of the ALARRM data obtained to date: 1035 observations spread over 165 target fields, representing approximately 100 days of telescope time with AMI-LA (approx. 100 GB of raw data). The calibration and imaging stages took around eight and six hours respectively on a single CPU-core. The metadata associated with these reductions were then analysed to confirm that the pipeline generally behaves as expected, identify any problematic datasets, and place some quantitative values on the benefits of applying the full **chimenea** algorithm, as opposed to simply applying a single Clean to each observation.⁸

The first step of the analysis was to exclude any fields of view for which no sources were detected in the concatenated deep-image at the 5.5 sigma level — these fields have only a single open Clean operation performed at each epoch, based on the RMS estimated in the dirty map, and so do not provide relevant datapoints for this comparison. This left 862 observations spread over 121 target fields where a source was detected in the deep image. The data were reduced with (somewhat arbitrarily chosen) re-Clean convergence criteria of a RMS-decrease less than 5% after applying a Clean operation, or a maximum of three re-Clean iterations (cf Figure 2). For the ALARRM dataset, these values appear quite suitable — 503 observations converged after one iteration (i.e. less than 5% difference in estimated RMS between dirty map and image produced from one masked Clean operation), 296 converged after two iterations, leaving 63 observations for which the re-Clean algorithm iterated the maximum of three times. Of these 63 remaining observations, 62 had an RMS decrease in the final re-Clean iteration of less than 5%. Therefore we can consider only one observation as requiring a ‘hard-stop’ on the re-Clean iterations. (This non-converging observation turns out to be from a field with extended emission, which is generally problematic for the current **AMISurvey** pipeline.)

As an additional check on the pipeline’s behaviour, we produced Figures 4 and 5, which compare the fractional decrease in RMS-estimate between the dirty map and the final image, and the flux-density of the brightest source detected in the deep image. Generally, we expect fields with brighter sources to exhibit higher levels of side-lobe artefacts in the dirty map, which positively bias the initial dirty-map RMS-estimate (particularly for short observations with poor *uv*-coverage), and so we expect some level of correlation between these two quantities. The resulting plots show a smooth progression in the maximum RMS-decrease with increasing source flux, excepting some well-understood outliers:

- The observations of XTEJ908+094 are affected by a very bright out-of-field source, which presents extended side-lobe artefacts as additional background variation in the field-of-view.
- The PTF09AXC field contains extended emission and/or blended sources (which are not well suited to reduction by the automated pipeline).
- The observations of field SWIFT_554620 contain a single epoch with a particularly bright source (GRB140327A in afterglow), and so this observation falls lower on the plot than the rest of the grouping.

In total there are 84 observations of fields with sources brighter than 10 mJy, which are not shown in Figures 4 and 5. When plotted, these datapoints continue the general trend with increased scatter, but we do not consider them as suitable for automated reduction in the current **AMISurvey** pipeline as modified calibration settings need to be applied when a source of around 10 mJy or brighter is in the field of view.

As can be seen in Figure 4, when a bright source is present the RMS-estimates may drop by a factor of two or more after the masked re-Clean process is applied (i.e. a proportional drop to less than 0.5 times the original value, in the units plotted). Since this RMS-estimate is then used when applying the open-Clean process to look for single-epoch transients, it suggests a potentially significant improvement in sensitivity to transient sources when a bright steady source is present in the field of view. However, some observations of fields with a bright steady source do not show such improvement. This is to be expected if short observations, or those which undergo a high degree of RFI flagging, have sufficiently poor data quality that bright source sidelobes are not the dominant contributing factor to background variation. To test this hypothesis, we made a simple cut on the data by replotting only observations of duration greater than 3.5 hours, as shown in Figure 5 (leaving 289 observations spread over 65 different targets). This plot shows a much clearer relation between bright-source flux-density and RMS-estimate reduction.

We note that the RMS reduction factors (between dirty-map RMS-estimate and final RMS-estimate) displayed in Figures 4 and 5 give an indication of the improved imaging quality of the **AMISurvey** pipeline compared to our early reductions, since the naively-reduced images would be cleaned to a threshold determined by the dirty-map estimated RMS. The addition of a masked-Clean process and output image should also help to avoid under-estimation of flux in bright sources (‘Clean-bias’). However, the ultimate figures-of-merit relevant to transient surveys are those relating to sensitivity and accuracy in transient detection. Experience suggests the trade-off between sensitivity to real transients and artefacts in radio-astronomy is difficult to quantify without extensive testing against both real and simulated visibility data, which we leave to future investigation.

⁸The IPython notebook used to perform this analysis is included with the **AMISurvey** package, and a static version relating to the dataset described herein may be viewed at <http://nbviewer.ipython.org/gist/timstaley/5ef6dfc2a370e0288bb0>.

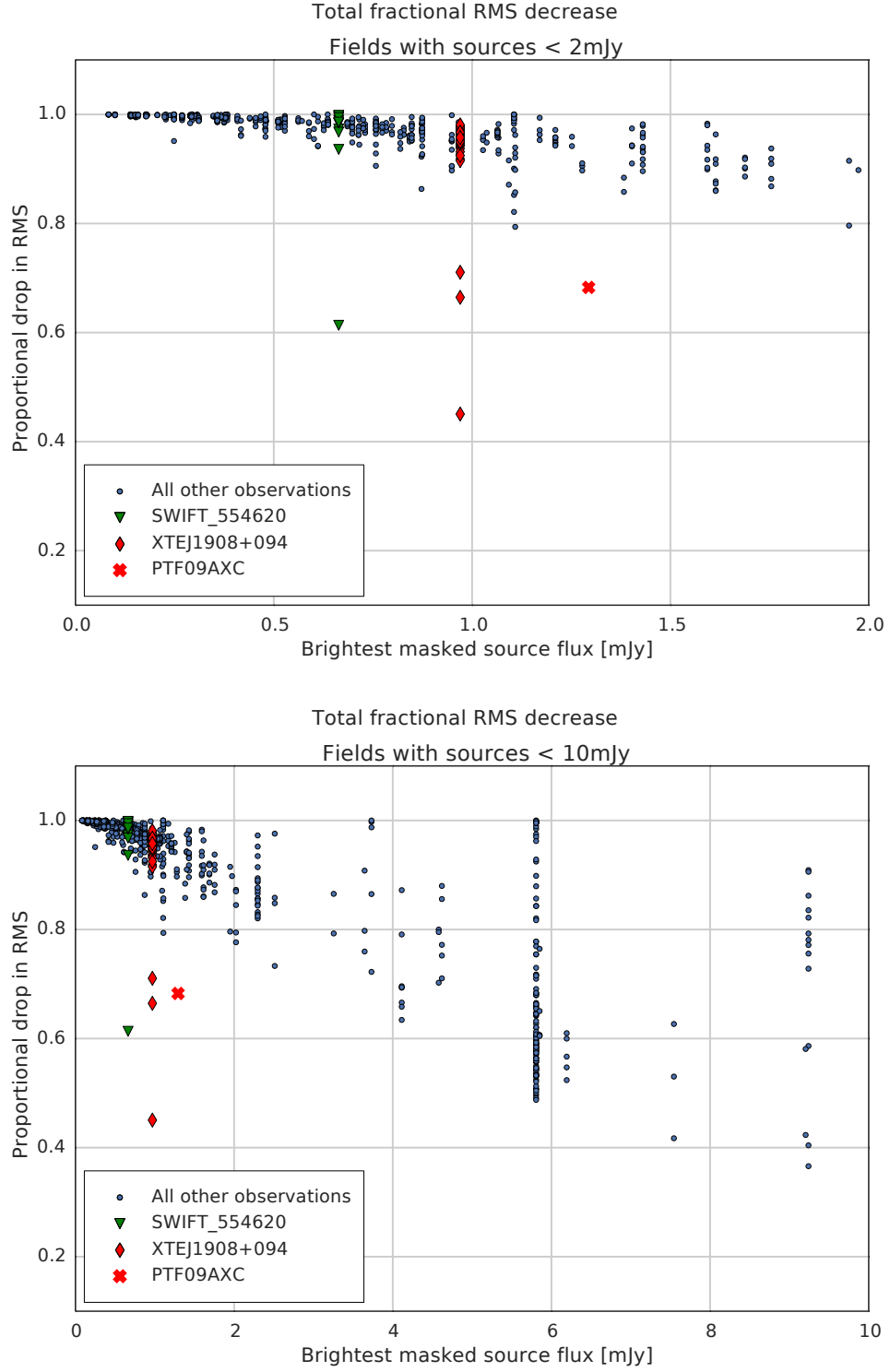


Figure 4: Scatter-plots comparing the fractional decrease in RMS-estimate (the ratio of [final image RMS / dirty map RMS] for each single-epoch observation), and the flux-density from the brightest source-detection in the deep image of that observation’s field of view. Each target has a single value for ‘brightest source-flux in deep image,’ but RMS-decrease varies between epochs, resulting in all RMS-decrease values for a single target being plotted as a column of points. As expected, fields with brighter sources in view often have more drastically reduced RMS levels than fields with only faint sources. A few outliers (overplotted in red and green) buck this trend, for well understood reasons — details in text. Upper and lower sub-plots depict the same data but with different limits on plotting region applied, to give more clarity in the densely populated upper-left region.

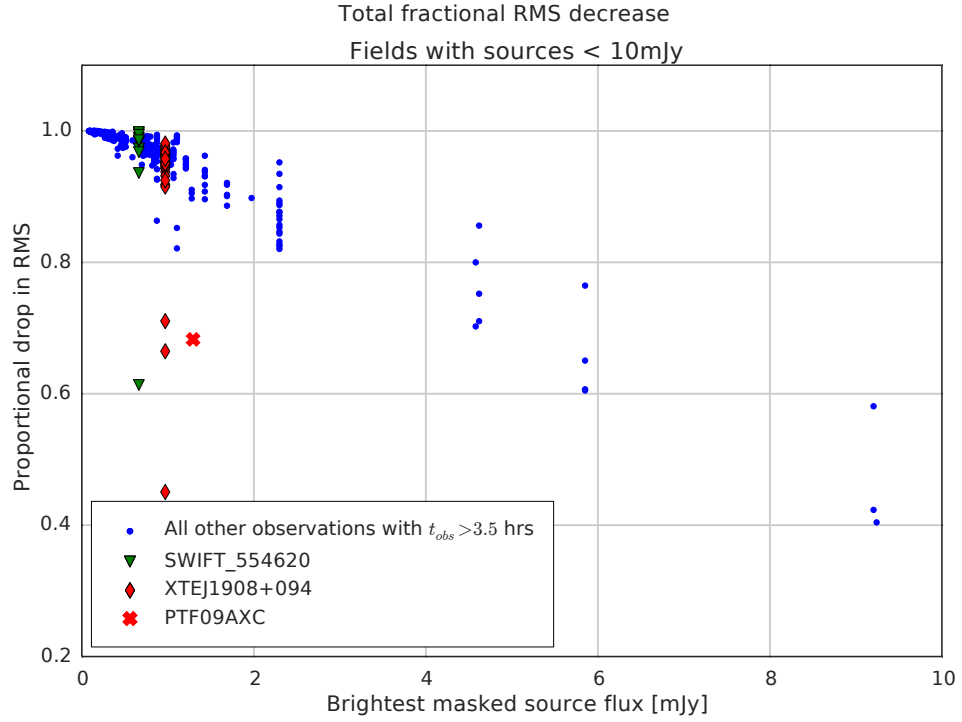
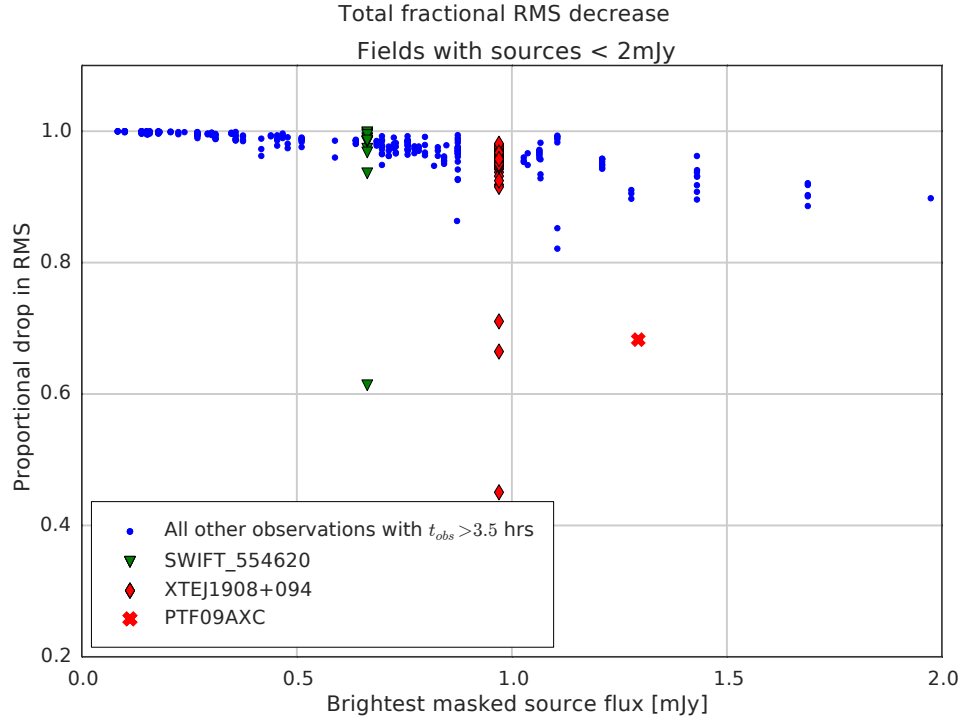


Figure 5: As Figure 4, but this time only observations of duration 3.5 hours or greater (typically 4 hours) are plotted. With the short-duration observations removed, the plots show a clearer relation between the flux-density level of the brightest source in the field, and the RMS-estimation improvement gained by applying an iterative Clean process.

7. Discussion

7.1. Limitations, possible extensions and alternatives to *chimenea*

While *chimenea* and the other packages presented here have proven useful in the reduction of ALARRM datasets, the approach is not without its limitations. The most immediate problem is that the RMS-estimation procedures do not work well in the presence of extended emission or complex blended sources, and so fields containing such sources still need to be manually reduced (though these types of source are always tricky when reducing radio-synthesis data). The iterative approach also adds a non-trivial amount of computational overhead, since a field of view may be processed with a Clean operation multiple times as the background RMS is re-estimated. The ALARRM datasets to date have been relatively quick to reduce, with file sizes on the order of a gigabyte or less and reduction times on the order of minutes, but for a large dataset the re-Cleaning may become a significant contribution to the overall processing time.

We should also reiterate that this work has been motivated to date by reproducing the data-reduction steps we would previously have applied manually. With this benchmark implementation in place, we can begin to consider alterations which might improve sensitivity and accuracy. A logical extension would be to utilise *uv*-plane model fitting, to improve the accuracy of flux-density measurements after candidate sources have been identified via the current pipeline. This avoids the degeneracies involved in ‘fitting a model to a model’ — i.e. fitting a source-model to a synthesized image that has already been produced by inference from the raw data (Martí-Vidal et al., 2014). Additionally, it would be interesting to compare the results from our pipeline to some of the more recent source-detection and modelling approaches that have been developed on the basis of Bayesian inference (e.g. Sutter et al., 2014; Lochner et al., 2015), since these provide a full modelling of the data from the *uv*-plane information, although a significant trade-off in terms of increased computational overhead is to be expected. Image-plane transient-detection pipelines for the SKA will have stringent demands on both accuracy and computational budget, and so it may be that some hybrid or simplified image-synthesis approach will turn out to be most attractive in the long run.

7.2. Methods for interfacing with legacy or proprietary software

The *drive-ami* and *drive-casa* software packages discussed in this paper solve a problem commonly found throughout software-intensive fields of both academia and industry — how to employ (in an automated fashion) a software package that cannot be directly called as a library. This may arise when the external package is written in a different programming language to that of the user’s codebase, or because the external codebase is legacy or

proprietary and only the command-line interface is accessible. In this subsection we describe three approaches that may be adopted when creating an interface to legacy or cross-language software, together with their merits and drawbacks, and explain why we chose to use terminal-emulation in this work.

7.2.1. Cross-language bindings

Typically, the core algorithms of data-reduction pipelines are implemented in a lower-level language such as Fortran, C or C++, either for performance reasons or simply because these languages were most suitable when the software was initially written. However, it is generally recognised that high-level programming languages such as Python, Perl or IDL provide a more suitable environment for rapid prototyping and development of scientific algorithms, due to their flexibility, succinctness, and convenient access to a large variety of external libraries for tasks such as data-visualization or statistical analysis. The standard approach to making routines written in a low-level language available from a higher-level language is to implement cross-language “bindings”, specialized modules which map data-structures and function calls between the two. Examples of this approach in radio astronomy are *miriad-python* (Williams et al., 2012), which provides Python bindings for *miriad* (Sault et al., 1995; Sault et al., 2011), and the *Obit* / *ParselTongue* packages (Cotton, 2008, 2013; Kettenis et al., 2006; Kettenis and Sipior, 2012), which provide Python bindings for AIPS (Fomalont, 1981). We refer the reader to Williams et al. (2012) for further discussion on these packages and their merits.

Cross-language binding modules are usually the most desirable solution for developing a scriptable interface to routines written in a low-level language, since a well-written set of modules will provide a high-performance, seamless interface between the external package and the user’s programming language of choice. However, there are associated costs and requirements that may make such an approach impractical. In astronomy, the primary issues with software development are usually those of available developer time and expertise. Astronomers with knowledge of both high-level and low-level languages, and how to implement bindings between them, are rare. Even if someone with the requisite skills is available, the task of creating bindings and porting the required pipeline-logic to a high-level language may require a significant investment of developer time, both to understand the logic implemented in the legacy system, and to verify that it is faithfully reproduced in the ‘ported’ version. At the end-user level, there is also the problem of familiarity: a cross-language port of a well-used legacy package will require users to re-learn the interface via its new bindings. Finally, there are the twin issues of access to source code, and maintenance — if a tool is proprietary, or if the source is unavailable for other reasons, or non-trivial to compile, then this presents an additional hurdle to implementation and use of bindings. Any changes to the original package must also be reflected

in the bindings modules, requiring code-maintenance and recompilation on a schedule tied to the external package.

As such, implementing binding modules from scratch may be overkill. Some alternative approaches to using low-level or legacy codes from a higher-level language are described below.

7.2.2. Sub-process calls

If an external tool is usually invoked via an executable which takes command line arguments and then requires no further interaction, it is often trivial to invoke in a scripted fashion (otherwise known as ‘spawning’ or ‘calling out’ to a sub-process). The built in Python library `subprocess` provides a perfectly adequate interface when this is the case. We also direct the reader towards the `sh` package⁹, which provides an alternative interface in which external commands may be represented by objects akin to Python functions, making for a neater syntax in repeated usage.

However, this approach falls down when invoking longer-running, multi-step processes that may require user interaction. While such tools often provide an option to run a series of commands from a script, parsing the concatenated output from many individual commands to record details of the data-reduction, or worse yet detecting error conditions and locating the cause, can present a considerable challenge. In addition, larger, interactive tools often incur a few seconds start-up time while the environment loads, and so repeatedly invoking the tool for multiple smaller steps can be impractically slow.

7.2.3. Terminal emulation

A third option is that of terminal emulation: spawn the external tool in a sub-process, but then interact with it by emulating interactive terminal input. While this is slightly more involved than a simple sub-process call it allows for a one-to-one mapping of the sub-routines presented by the external tool into the user’s scripting language of choice, without any of the source modification or recompilation that implementing bindings requires.

In Python, terminal emulation functionality can be provided by a package called `pexpect`¹⁰. The `pexpect` package file is small (a 132kB download) and only requires a working Python installation (v2.6 or greater) as a prerequisite.

When building an interface to a legacy tool using `pexpect` in Python, we have found that it usually makes sense to provide an interface class. This class can be used to spawn the external process when initialized, store a reference to the `pexpect` object representing that process and the emulated terminal interface, handle the terminal interaction, and so on. Implementation of the basic interface class is relatively simple. The essential parameters are a regular expression representing the format of the command-line prompt presented by the external tool, the path to

invoke the external tool, and any required environment variables. Once the underlying interface is in place, end-user scripts that would previously be invoked from within the external tool’s interactive environment can immediately be employed as part of a larger pipeline. Furthermore, the precise inputs passed via the interface can be recorded alongside any data-outputs, providing data-provenance, so that if an end-user later wishes to inspect the data-reduction process manually, they can load up the external tool and step through the recorded command-script, either verifying or varying the reduction steps as required.

From this starting point, it is straightforward to build up a more comprehensive interface suitable for use in complex logic flows. Since the output from each command is available in a clearly separated block of text, it is often trivial to implement basic parsing routines that can recover information from the command-line output of the external tool, allowing for ‘if / else’ logic branches in the larger pipeline which make use of intermediate outputs from the external tool. If desirable, convenience routines can be added to ease the process of generating the commands passed to the external tool, providing an interface which is better suited to the high-level scripting language (while still allowing direct access to the external tool as required).

Since this approach effectively treats an external tool’s command-line interface as a programming interface, any updates to the external tool are unlikely to require changes to the scripting-interface, unless they involve changes that also change the user-interaction behaviour (with one caveat — if output-parsing relies on a particular output format, this may be changed to another layout which is still human-friendly but breaks the fixed parser rules).

In terms of performance, terminal emulation sits somewhere between that of cross-language bindings and simple sub-process invocation. Small delays (typically around 0.05 seconds) in the input-output cycle must be allowed for when emulating terminal interaction, presumably due to the small but measurable time it takes for the external tool to parse the commands input. In addition, any overhead incurred when starting the external tool is still present. However, once initialized, an instance of the external tool can be re-used for many commands, and so this start-up overhead is incurred fewer times when compared to simple sub-process calls.

It is this terminal-emulation approach that we adopted in implementing Python interfaces to the AMI-REDUCE and CASA packages, for reasons outlined below.

7.2.4. Use of terminal-emulation in *drive-ami* and *drive-casa*

The AMI-REDUCE pipeline is implemented in FORTRAN 77, and comprises around 58,000 lines of source-code¹¹. In the rare case that AMI-REDUCE needs to be run elsewhere

⁹ <http://amoffat.github.io/sh/>

¹⁰ <https://github.com/pexpect/pexpect>

¹¹ As analyzed using the SLOCCount utility:
<http://www.dwheeler.com/sloccount/>
(Including component libraries.)

than the MRAO it is typically distributed as a pre-built binary, with a copy of the source files included. Despite the author's efforts to install appropriate dependencies and alter makefiles, building the package from source on a modern Ubuntu Linux installation proved to be problematic (as opposed to simply executing the pre-built binary, which was simple once the appropriate compatibility packages had been downloaded and installed). Given the size and complexity of the codebase, the difficulties building from source, and the authors' lack of expertise in Fortran, we judged that implementing Python-bindings to the AMI-REDUCE FORTRAN routines directly would require a significant and costly investment of developer time, with little benefit given that AMI-REDUCE is not a widely-used software package. As described in Section 2, simple sub-process calls to AMI-REDUCE were initially used in our end-to-end pipeline, but soon proved inadequate. Terminal-emulation presented a compromise - greater control and flexibility in interfacing with AMI-REDUCE, without requiring a large effort to create FORTRAN/Python binding modules. The added logging capabilities also provide reassurance that our data-reduction process is performing as expected, and can be verified manually if required. Development of the `drive-ami` interface using terminal emulation allowed for a relatively low initial time-investment to create a basic interface class. Additional features such as error-handling and metadata-extraction were then added piecemeal on an 'as needed' basis, since addition of minor features can typically be accomplished in a few hours.

The case for writing a terminal-emulation tool for CASA is less obvious, since `casapy` already provides a scriptable Python interface — on the face of it, writing a Python library to interface with a Python interpreter is faintly ridiculous. However, as detailed in Section 3, use of CASA routines alongside other Python libraries is otherwise a difficult and error-prone process. In the long-term this may change, but for now the additional logging capabilities, together with the fact that `drive-casa` can be trivially installed on most systems, mean that it provides significant benefits to the end-user in terms of both convenience and reproducibility.

7.3. Writing re-usable Python codes

This work highlights a pressing issue for the field of astronomy as a whole (and wider academia): that of software interoperability and re-use. There will always be a trade-off to be made in choosing to incorporate legacy packages into new efforts, or starting afresh with a re-write in the language *du jour* — nothing has changed there. However, when writing new software, we would encourage authors to consider how their code might be adapted for re-use, either in part or in whole. In the wider context of software development, situations are so varied that it is hard to give any firm advice on this topic (or at least, beyond the scope of this work). However, if we restrict our scope to that of writing scientific codes in Python, we can be a little

more concrete; giving a few examples from this work that hopefully serve to illustrate the wider argument:

- Prefer interoperable, modular libraries over monolithic pipeline packages (this is a widely discussed viewpoint more generally known as the *Unix philosophy*). The packages described in this paper began life as one big collection of prototype code that called out to each external tool in turn (via a sub-process call). As we added features and the codebase grew, it became clear that it was worth separating into separate packages.
- Use the Python packaging infrastructure. A major hurdle to breaking up a code into sub-packages is that it induces additional dependency management. Embracing the Python packaging infrastructure to implement a versioned install process helps keep this under control, and also makes it much easier to distribute the code to users not directly involved in the development process.¹²
- Ideally, distinguish technical implementation details from higher-level applications of code (or 'low-level' vs 'high-level' programming interfaces). Core libraries should ideally be simple (each function does one thing well) and flexible (functions can be used independently, perhaps requiring a number of parameters, etc). However, when writing a larger software package it is essential to hide common-usage patterns of a low-level interface behind larger subroutines - this is the only way to write complex code which is readable and maintainable. Think carefully about where it might be sensible to divide these code hierarchies and separate them into high-level and low-level packages. For example, in this work `drive-casa` provides a direct interface to the `casapy` package (low-level), while `chimenea` encodes the scientifically relevant logic and implements a particular view of how best to structure the data (high-level).
- Avoid use of the `print` function. The Python `logging` package provides a good degree of control and flexibility in how debugging and information messages are displayed, stored to file, distributed over the network, etc, whereas the `print` function is a basic on/off output and should only ever be used as a temporary measure — otherwise you risk overwhelming calling routines with a flood of irrelevant debugging output. There are also many extensions to the `logging` functionality, for example `AMISurvey` makes use of the `colorlog` package¹³ to colourise the terminal output by logging category.

¹²A guide to creating installable Python packages may be found at <http://the-hitchhikers-guide-to-packaging.readthedocs.org/>

¹³<https://github.com/borntyping/python-colorlog>

- Separate configuration from code. Ideally, any scientifically significant parameters should be clearly identifiable and easily changed from one location. In the current `AMIsurvey` pipeline, telescope and data specific parameters are encoded in a single, short module dedicated to the AMI-LA telescope. This could easily be extended to load a different module depending on metadata entries, or load parameters from a user-edited text file as required.

Finally, we highly recommended that even if a scientific code follows none of this advice, it should still be made freely available. Stylistic changes to suit a given context can always be made after release, and the benefits of source code sharing are significant, perhaps most crucially for encouraging reproducibility (see e.g. Shamir et al., 2013, for further discussion).

8. Conclusions

In this work we have described an automated pipeline for reduction of radio-synthesis data obtained for transient follow-up, and undertaken initial performance tests which show it behaves as desired. Although heuristic in approach, `chimenea` already provides a useful component in automatic data-reduction pipelines, and serves as a test-bed and benchmark for more sophisticated reduction algorithms. We have highlighted limitations to the pipeline and discussed possible avenues for future development and exploration. We note that extensive further development in the area of automated imaging will be essential for performing real-time image-plane transient surveys, one of the key science drivers for the SKA (Burlon et al., 2015; Corbel et al., 2015; Pérez-Torres et al., 2014).

From a technical standpoint, we have demonstrated the effectiveness of terminal-emulation as a method for interfacing with legacy or externally maintained software. We found this approach relatively easy to implement and well suited to our needs, especially given the added benefits of easy manual reproducibility and data-provenance.

Finally, we have provided some basic guidelines and examples towards writing re-usable Python codes in astronomy. This topic could be the subject of a much larger work, but we hope the suggestions herein will stimulate further thought and discussion. In particular, we hope that our efforts to separate and document the components of the `AMIsurvey` pipeline will result in productive re-use elsewhere.

Acknowledgements

We thank the staff of the Mullard Radio Astronomy Observatory for their invaluable assistance in the operation of AMI.

This work made use of a number of publicly available software packages and data products: NASA’s Astrophysics Data System; Astropy, a community-developed core Python

package for Astronomy (Astropy Collaboration et al., 2013); the IPython package (Pérez and Granger, 2007); matplotlib, a Python library for publication quality graphics (Hunter, 2007); and SciPy (Jones et al., 2001).

This project was wholly funded by European Research Council Advanced Grant 267697 ‘4 π sky: Extreme Astrophysics with Revolutionary Radio Telescopes.’

References

References

- Anderson, G. E. et al., May 2014. Probing the bright radio flare and afterglow of GRB 130427A with the Arcminute Microkelvin Imager. *Mon. Not. R. Astron. Soc.* 440, 2059–2065.
- Astropy Collaboration et al., Oct. 2013. Astropy: A community Python package for astronomy. *Astron. Astrophys.* 558, A33.
- Bell, M. E., et al., Feb. 2014. A survey for transients and variables with the Murchison Widefield Array 32-tile prototype at 154 MHz. *Mon. Not. R. Astron. Soc.* 438, 352–367.
- Booth, R. S., & Jonas, J. L., Mar. 2012. An Overview of the MeerKAT Project. *Afr. Skies* 16, 101.
- Burlon, D., Ghirlanda, G., van der Horst, A., Murphy, T., Wijers, R., Gaensler, B., Ghisellini, G., & Prandoni, I., Jan. 2015. The SKA view of Gamma-ray Bursts. *ArXiv e-prints*.
- CASA Consortium, Jul. 2011. CASA: Common Astronomy Software Applications. *Astrophys. Source Code Lib.*, record ascl:1107.013.
- Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., & Broderick, J. J., May 1998. The NRAO VLA Sky Survey. *Astron. J.* 115, 1693–1716.
- Corbel, S. et al., Jan. 2015. Incoherent transient radio emission from stellar-mass compact objects in the SKA era. *ArXiv e-prints*.
- Cotton, B., Jul. 2013. Obit: Radio Astronomy Data Handling. *Astrophys. Source Code Lib.*, record ascl:1307.008. URL <http://ascl.net/1307.008>
- Cotton, W. D., Apr. 2008. Obit: A Development Environment for Astronomical Algorithms. *Publ. Astron. Soc. Pac.* 120, 439–448.
- Davies, M. L. et al., Dec. 2009. Follow-up observations at 16 and 33GHz of extragalactic sources from WMAP 3-yr data: I - Spectral properties. *Mon. Not. R. Astron. Soc.* 400, 984–994.
- Fender, R. P., Anderson, G. E., Osten, R., Staley, T., Rumsey, C., Grainge, K., & Saunders, R. D. E., Jan. 2015. A prompt radio transient associated with a gamma-ray superflare from the young M dwarf binary DG CVn. *Mon. Not. R. Astron. Soc.* 446, L66–L70.
- Fomalont, E., 1981. Astronomical Image Processing System / AIPS. *Natl. Radio Astron. Obs. Newsl.* 3, 3.
- Högbom, J. A., Jun. 1974. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *Astron. Astrophys. Suppl.* 15, 417.
- Hunter, J. D., 2007. Matplotlib: A 2d graphics environment. *Comput. Sci. Eng.* 9 (3), 90–95.
- Jones, E., Oliphant, T., Peterson, P., & Others, 2001. SciPy: Open source scientific tools for python. URL <http://www.scipy.org/>
- Kettenis, M., & Sipior, M., Aug. 2012. ParselTongue: AIPS Python Interface. *Astrophys. Source Code Lib.*, record ascl:1208.020. URL <http://ascl.net/1208.020>
- Kettenis, M., van Langevelde, H. J., Reynolds, C., & Cotton, B., Jul. 2006. ParselTongue: AIPS Talking Python. In: Gabriel, C., Arviset, C., Ponz, D., & Enrique, S. (Eds.), *Astron. Data Anal. Software and Syst.* XV. Vol. 351 of Astronomical Society of the Pacific Conference Series. p. 497.
- Lochner, M., Natarajan, I., Zwart, J. T. L., Smirnov, O., Bassett, B. A., Oozeer, N., & Kunz, M., Jun. 2015. Bayesian inference for radio observations. *Mon. Not. R. Astron. Soc.* 450, 1308–1319.
- Martí-Vidal, I., Vlemmings, W. H. T., Muller, S., & Casey, S., Mar. 2014. UVMULTIFIT: A versatile tool for fitting astronomical radio interferometric data. *Astron. Astrophys.* 563, A136.

- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K., Oct. 2007. CASA Architecture and Applications. In: Shaw, R. A., Hill, F., & Bell, D. J. (Eds.), *Astron. Data Anal. Software and Syst.* XVI. Vol. 376 of Astronomical Society of the Pacific Conference Series. p. 127.
- Metzger, B. D., Williams, P. K. G., & Berger, E., Jun. 2015. Extragalactic Synchrotron Transients in the Era of Wide-field Radio Surveys. I. Detection Rates and Light Curve Characteristics. *Astrophys. J.* 806, 224.
- Miller, N. A., Fomalont, E. B., Kellermann, K. I., Mainieri, V., Norman, C., Padovani, P., Rosati, P., & Tozzi, P., Nov. 2008. The VLA 1.4 GHz Survey of the Extended Chandra Deep Field-South: First Data Release. *Astrophys. J.* 179, 114–123.
- Mooley, K. P., Frail, D. A., Ofek, E. O., Miller, N. A., Kulkarni, S. R., & Horesh, A., May 2013. Sensitive Search for Radio Variables and Transients in the Extended Chandra Deep Field South. *Astrophys. J.* 768, 165.
- Murphy, T., et al., Feb. 2013. VAST: An ASKAP Survey for Variables and Slow Transients. *Publ. Astron. Soc. Aust.* 30, 6.
- Norris, R. P. et al., Mar. 2013. Radio Continuum Surveys with Square Kilometre Array Pathfinders. *Publ. Astron. Soc. Pac.* 30, 20.
- Pérez, F., & Granger, B. E., May 2007. IPython: a system for interactive scientific computing. *Computing in Science and Engineering* 9 (3), 21–29.
URL <http://ipython.org>
- Pérez-Torres, M. A. et al., Sep. 2014. Core-collapse and Type Ia supernovae with the SKA. ArXiv e-prints.
- Ridgway, S. T., Matheson, T., Mighell, K. J., Olsen, K. A., & Howell, S. B., Nov. 2014. The Variable Sky of Deep Synoptic Surveys. *Astrophys. J.* 796, 53.
- Sault, R. J., Teuben, P. J., & Wright, M. C. H., Jun. 2011. MIRIAD: Multi-channel Image Reconstruction, Image Analysis, and Display. *Astrophys. Source Code Lib.*, record ascl:1106.007.
URL <http://ascl.net/1106.007>
- Sault, R. J., Teuben, P. J., Wright, M. C. H., Payne, H. E., & Hayes, J. J. E., 1995. A retrospective view of miriad. In: Shaw, R. A. (Ed.), *Astron. Data Anal. Software and Syst.* IV. Vol. 77. p. 433.
- Schwab, F. R., Jul. 1984. Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry. *Astron. J.* 89, 1076–1081.
- Shamir, L. et al., Feb. 2013. Practices in source code sharing in astrophysics. *Astron. Comput.* 1, 54–58.
- Staley, T. D., Apr. 2015. drive-casa: Python interface for CASA scripting. *Astrophys. Source Code Lib.*, record ascl:1504.006.
- Staley, T. D., & Anderson, G. E., Feb. 2015a. AMIsurvey: Calibration and imaging pipeline for radio data. *Astrophys. Source Code Lib.*, record ascl:1502.017.
- Staley, T. D., & Anderson, G. E., Apr. 2015b. chimenea: Multi-epoch radio-synthesis data imaging. *Astrophys. Source Code Lib.*, record ascl:1504.005.
- Staley, T. D. et al., Feb. 2013. Automated rapid follow-up of Swift gamma-ray burst alerts at 15 GHz with the AMI Large Array. *Mon. Not. R. Astron. Soc.* 428, 3114–3120.
- Stappers, B. W., Apr. 2013. The square kilometre array and the transient universe. *Philos. Trans. R. Soc. London, Ser. A* 371, 20284.
- Sutter, P. M. et al., Feb. 2014. Probabilistic image reconstruction for radio interferometers. *Mon. Not. R. Astron. Soc.* 438, 768–778.
- Swinbank, J. D. et al., Jun. 2015. The LOFAR Transients Pipeline. *Astron. Comput.* 11, 25–48.
- Tingay, S. J. et al., Jan. 2013. The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies. *Publ. Astron. Soc. Aust.* 30, 7.
- TraP contributors, Dec. 2014. TraP: Transients discovery pipeline for image-plane surveys. *Astrophys. Source Code Lib.*, record ascl:1412.011.
- van Haarlem, M. P. et al., Aug. 2013. LOFAR: The LOW-Frequency ARray. *Astron. Astrophys.* 556, A2.
- White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D., Feb. 1997. A Catalog of 1.4 GHz Radio Sources from the FIRST Survey. *Astrophys. J.* 475, 479–493.
- Williams, P. K. G., Law, C. J., & Bower, G. C., Jun. 2012. Rapid Development of Interferometric Software Using MIRIAD and Python. *Publ. Astron. Soc. Pac.* 124, 624–636.
- Zwart, J. T. L. et al., Dec. 2008. The Arcminute Microkelvin Imager. *Mon. Not. R. Astron. Soc.* 391, 1545–1558.